

Review of
“Racial Fairness in Sentencing: A Case Study of Selected Crimes in Jefferson County”

Review by Brian Ostrom, Ph.D.
Senior Research Associate, National Center for State Courts
July 8, 2005

- **Purpose of Study** – the purpose of the study is to examine racial fairness in sentencing during the 1999-2002 period for cocaine possessing and misdemeanor shoplifting. While the goal of studying racial fairness is clear, it is not clear why the authors chose to focus on these two offense types. It would seem that the place to start (or at least include) is some type of crime against person (e.g., robbery, assault, criminal sexual conduct). The water surrounding drug possession offenses is quite muddy as the crack epidemic has receded and attitudes toward drug offenses have been changing (e.g., dropping of mandatory minimums). While this may not be the case in Kentucky, the possibility exists. For this reason, I would suggest focusing on a different class of offenses.
- **Methodology** – the authors of the study apply the “best practice” methodology in the field of sentencing research. In principle, the authors have attacked the problem in the right way. What I find especially innovative is their attempt to apportion the differences between black and white offenders into differences due to endowments and differences due to discrimination. It is clear that the authors know how to set up a methodologically appropriate study. In addition, the authors do an outstanding job of discussing the problems associated with sample selection bias – offenders are systematically being excluded from the sample of convicted offenders through the differential treatment that can occur at detection, charging, and conviction stages of the process.
- **Data** – From the evidence supplied in the report, it appears that there are some potentially severe missing data problems: Cocaine – have complete data on only 825 of 1,035 cases; Shoplifting – have complete data on 1,599 of 4,232 cases. What accounts for the lack of data? Is there anything systematic about the missing data? All in all, this raises a red flag. While the authors are clear about sample selection bias, there is ample reason to be concerned about the failure to collect a complete data set. I know that getting complete data on all cases is always challenging. However, according to the report, most of the missing data is due to a lack of information on use of public defender and judge. It would seem that some effort should have been directed toward overcoming these omissions.
- **Sentencing Models** – According to the report, the authors had data on 19 variables plus individual judges dummy variables. An examination of Table 2 suggests that they have a sufficiently rich data set.
 - I would suggest that the dummy variable for age be reconsidered. There is work in the field that suggests that younger offenders are treated differently than those in their mid-twenties. Also, the possibility exists of using some interaction variables like Steffensmeier et al; for example, looking at young, black, males.

- A problem arises when the actual estimated models are presented as they only include variables that are statistically significant at least once in one of the equations. As a modeling strategy, we find this to be somewhat haphazard. This is especially true in light of the authors' finding that the coefficient of one insignificant variable in Table 9 is viewed as being important. Given issues of collinearity, I think that some a priori decision about the variables that ought to influence a sentencing decision should be included in the model.
 - Given the paucity of variables in the analysis, I am somewhat suspicious of the inclusion of the individual judge variables in the in/out models. A significant value for one of these variables could be drastically affected by omitted variable bias for example.
 - We do not know the mean values for the judge variables; we do not know what roles the judges have (arraignment, coaxing plea agreements, etc). Also given the fact that there is missing data because of judge name, it seems wise to exclude these variables from the model.
 - We have some concern that important variables that are taken into consideration by judges are omitted prematurely. This is borne out in the shoplifting model when less than 5% of the total variance is explained
 - Question about the authors' interpretation of the public attorney variable – it would be nice to know a little bit more about the indigent defense system in KY. I think it would be wise to pursue some alternative functional forms to see if the finding is robust to other model specifications.
- **Application of Methodology** – on the whole, the application of the methodology is first rate. The only problem that I see is that the hazard rate should not have a negative sign – footnote 23 is in error based upon our understanding of the concept. We really appreciate the insights offered by the Blinder-Oaxaca decomposition and would encourage the authors to keep this as part of their research arsenal.
 - **Findings** – the authors do not do much with the estimated model to make it possible to gain some substantive insights into the sentencing process. For example, calculating the impact of each of the variables on the probability of incarceration holding remaining variables at the mean would provide a way to determine if the statistically significant race variable (where it occurs) makes a substantive difference.
 - **Significance of Findings**
 - The percent correctly predicted in Tables 6, 7, and 9 needs to be compared to the null model of Black 68%, Other 82%, and Pooled 72% for cocaine model and 90% Black, 95% Other, and 93% Pooled for shoplifting model-- with this comparison the models do not do a very good job of improving over the null
 - We would not base any policy changes on the results of this study
 - Perhaps the primary use of the study is to introduce the court to empirical study and stress the importance of having as wide ranging and complete a data set as possible

- While there is a statistically significant difference in the cocaine in/out decision, there are enough concerns about data and methods to question the robustness of the conclusion
- **Recommendations** – we recommend the following
 - Expand the study to include violent offenses
 - Do what it takes to assemble a comprehensive and complete data set
 - Focus on the establishment of a substantively sound statistical model
 - Spend time to make the research design, findings and implications readily understandable and interpretable by a nontechnical audience. As written, the report is quite difficult to decipher by individuals not well versed in statistical theory.

On the whole, this is a competent application of the appropriate methodology. The presence of a large amount of missing data coupled with a narrow focus on only two offenses types, limits the generalizability of the present analysis. In addition, it is important to include more variables in the statistical model especially if one is interested in the impact of individual judges. Finally, the report needs to be written for an audience of judges, court administrators and legislative analysts. Reducing the level of technical jargon and increasing the discussion around the substantive policy significance would make the report more relevant for the intended audience.